

DOCUMENT RESUME

ED 188 911

SE 031 333

AUTHOR van den Berg, Euwe: And Others
TITLE The Convergent Validity of the Cognitive Preference Construct. Draft Copy.
PUB DATE Apr 80
NOTE 21p.: Paper presented at the annual meeting of the National Association for Research in Science Teaching (53rd, Boston, MA, April 11-13, 1980).
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Measurement; *Cognitive Processes; Cognitive Style; Critical Thinking; Educational Research; Higher Education; Learning Theories; *Science Education; *Validity
IDENTIFIERS Science Cognitive Preference Inventory; *Science Education Research

ABSTRACT

Reported is a study designed to investigate the convergent validity of the cognitive preference construct and to develop new techniques for analyzing cognitive preference data. The cognitive preference construct was invented to help evaluate the science curricula of the sixties based on the assumption that a student's intellectual approach to information was more important than ability to identify correct information. Four categories of intellectual approach as recall (R), principles (P), questioning (Q), and applications (A) are presented. Two new instruments to measure cognitive preferences were constructed and administered and a conventional instrument (the Science Cognitive Preference Inventory) was given to 71 elementary education majors. The results were analyzed using a multitrait-multimethod design. Developed were two statistical indices of distinctness of cognitive preference, one for ranking data and one for normative data. Correlations of Q and R with R, P, Q, and A scores across and within instruments, fulfilled most of the conditions for convergent validity. Correlations of P and A with R, P, Q, and A scores only partially fulfilled these conditions indicating weaknesses in the P and A modes. Correlations of Q and R strongly supported the interpretation of Q and P as opposite poles of one dimension which could be called "critical thinking."
(Author/GS)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED188911

Draft copy

The Convergent Validity of the Cognitive

Preference Construct

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Vincent Lunetta

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

by

Euwe van den Berg
c/o Office of International Relations
Free University
P.O. Box 7161
1007 MC AMSTERDAM
The Netherlands

Vincent N. Lunetta
Science Education Center
The University of Iowa
IOWA CITY, Iowa 52242
U. S. A.

Pinchas Tamir
Israel Science Teaching Center
The Hebrew University
JERUSALEM
Israel

Paper presented at the 53rd Annual Meeting of the National Association
for Research in Science Teaching, Boston, Massachusetts, April 11-13, 1980

SE 031 333

THE CONVERGENT VALIDITY OF THE COGNITIVE PREFERENCE CONSTRUCT

Introduction

The construct of cognitive preference (c.p.) was invented to serve in the evaluation of the science curricula of the sixties (Heath, 1964), based on the assumption that in the new curricula a student's intellectual approach to information would be more important than his/her "ability to identify correct or incorrect information". Heath then postulated preferences for four kinds of intellectual responses to information (adapted from Heath, 1964 and Tamir, 1975):

Factual information or recall (R): A preference for recall indicates an interest in learning a name, number, definition, formula, or fact without much processing of the information such as searching for relationships or attempting to assess the relevancy of facts.

Principles (P): A preference for principles indicates an interest in identifying and explaining (in terms of general theories or concepts) scientific rules and relationships.

Questioning (Q): A preference for questioning indicates an interest in critically analyzing and commenting on the validity of scientific information and in generating suggestions and hypotheses for further research.

Applications (A): A preference for applications indicates an interest in the application of scientific information.

Following Heath's (1964) successful evaluation study of the PSSC Physics Course the c.p. construct was used in many other evaluation studies of major secondary science curricula and their foreign adaptations (Tamir, 1977). Furthermore studies were conducted to

investigate relationships between c.p. , other cognitive styles, various measures of student achievement, student career interests, and teaching strategies. However, little attention was paid to the refinement and validation of the interpretation of the c.p. construct.

In a thorough article Brown (1975) pointed to several serious weaknesses in c.p. research and particularly in the measurement and interpretation of c.p. She suggested that the between-student variance on c.p. inventories (all of which had been constructed according to the c.p. measurement methodology developed by Heath) could be explained by other factors than cognitive preference, for example by student preferences for linguistic features, or by level of cognitive functioning. She also pointed out that no direct links had been established between c.p. and actual student cognitive activities in response to scientific information. Furthermore, Brown questioned the appropriateness of statistical methods and the ipsative nature of the c.p. construct. Some of Brown's criticism was shown to be invalid in studies by Tamir (1977), Tamir and Lunetta (1977), and Jungwirth (1978, 1979a, 1979b, and 1980), however, new questions were raised as well, particularly by Jungwirth. The basic problem appears to be conflicts between the c.p. construct as interpreted (often intuitively) by researchers and the measurement methodology of c.p. which seems to allow for alternative interpretations: construct validity is at stake.

One way of investigating construct validity and improving on the definition, interpretation, and measurement of a construct is to develop several mutually independent methodologies to measure the construct and to study the correlation between these methodologies. The objectives of this study were:

- 1) To develop several methodologies to assess c.p., one of which to be based on actual student cognitive behavior in learning situations;

- 2) To study the correlations between these methodologies in order to assess the convergent validity of the c.p. construct.

Construct Validity

According to Cronbach and Meehl (1955) a construct is some postulated attribute of people which has "certain associated meanings carried in statements of this general character: Persons who possess this attribute will, in situation X act in manner Y (with stated probability)". Construct validation entails the derivation of such statements from a theory involving the construct and the testing of these statements or hypotheses.

In his original definition of c.p. Heath already indicated some of the associated meanings of c.p. when he wrote: "When told that density is defined as mass per unit volume, one student may notice the utility of the density concept for identifying unknown substances, another may wonder about the effects of temperature on density change, still another might focus upon the equivalent notational expression, $D = M/V$. These different modes of attending to the subject matter will be called cognitive preferences". Note that this definition refers to spontaneous student responses to scientific information.

In the measurement methodology designed by Heath students are presented with test items consisting of a stem with some introductory information followed by four statements which each constitute a correct continuation of the stem and which represent one of the four modes of c.p. Students are then asked to rank or rate these statements on a 1-4 (ranking) or 1-6 (rating) point scale according to their liking or interest. The following is an example from the Science Cognitive Preference Inventory used in this study:

According to the atomic model of matter, the world is made up of small particles called atoms. Atoms are built from protons, neutrons, and electrons.

- A. The atomic model has guided scientists and engineers to many discoveries.
- B. Physicists are searching for particles that are possibly more fundamental than protons, neutrons, and electrons.
- C. Over one hundred different atoms have been identified.
- D. The number of protons/electrons in an atom determines how this atom will behave chemically when it comes close to other atoms.

Statement A refers to applications of the atomic model (A-mode of c.p.), B indicates that our understanding of the atomic model is not complete and that there are questions to be investigated (Q), C states a fact (R), and D is an expression of a principle with wide ranging implications (P). Note that this test measures cognitive preferences for explicitly stated information while Heath's original definition of c.p. at the beginning of this section referred to spontaneous student intellectual responses to information. However, it are these spontaneous responses educators are most interested in and interpretations of c.p. research have referred to frequently. Furthermore the test format is rather unusual to students since a c.p. inventory is a science subject matter test without "best" answers.

The validity of the c.p. construct and of c.p. inventories would be greatly enhanced if they could be linked with student spontaneous intellectual responses to scientific information in realistic learning situations. For example: "Would high Q (questioning) students ask more critical questions in class than low Q or perhaps high R (recall) students?", "Would high P (principles) students have a greater tendency to identify or deduce general rules or principles in a science text than a low P or high R person?", "Would a high A (applications) student be more interested in technology oriented readings than a low A person?".

Past research has established some indirect links between relevant student learning preferences and c.p. inventory R, P, Q, and A scores. Tamir and Kempa (1976) found that physical science majors in college

(Canada and Israel) were highest in P and lowest in A while engineering majors in Canada and medical students in Israel had, relative to other groups, the lowest scores in P and the highest in A, which is consistent with the interests of these groups (basic and applied science respectively). High school students who enroll in summer research programs could be expected to have a high preference for questioning and a low preference for recall, which was confirmed by Lunetta and Tamir (1978) and replicated with a different c.p. inventory by Van den Berg (1978). Tamir (1977) also found that medical students with an interest in teaching and research had higher Q-R scores (this "difference" score is usually being interpreted as a measure of scientific curiosity and a critical mind) than medical students with a primary interest in clinical work. In short, the studies quoted confirmed some hypotheses regarding the c.p. construct and actual learning/study preferences of groups of students.

Methodology

In this study an attempt was made to observe manifestations of c.p. in learning behavior of individual students. After some pilot studies it was decided to use two instruments in addition to a conventional c.p. inventory. One instrument was to assess student c.p.'s for explicit statements embedded in a popular science-like reading assignment, while the other instrument relied on teachers' assessments of spontaneous student classroom behavior thought to be related to c.p.

The data collected with these instruments were to be analyzed in the framework of the multitrait-multimethod methodology described by Campbell and Fiske (1959) and to be explained as follows: Each operational definition of a construct is a trait (construct) - method unit, a union of a particular trait content (for example attitude

toward science) with measurement procedures not specific to that content (a 60 item Likert scale). The trait or construct could be measured by other methods as well (interview, teacher rating) while the measurement method could be used for many other purposes (attitudes toward school, friends, etc.). The systematic variance among test scores may be due to individual differences between students in their reactions to the measurement features (the Likert scale) as well as to individual differences with respect to the trait (attitudes toward science). A way to estimate the "true" trait variance is to measure the trait using several independent methods ("multimethod", for example a Likert scale and an interview) and to correlate student scores on each of them. Such correlation coefficients have been known as validity coefficients.

However, measuring one trait using different methods still does not eliminate the possibility of confounding traits (one may be measuring attitude toward science teachers rather than attitude toward science). Therefore, in order to investigate the validity of one trait and its measurement one may measure some related traits, for each trait using two or more measurement methods ("multitrait-multimethod"). Trait validity then requires convergence of each trait across methods and discrimination between different traits across and within methods.

For cognitive preference, a 4-trait construct, construct validity requires amongst others (these are not the only criteria for construct validity as construct validity requires a comprehensive review of all research concerning the construct):

- 1) Corresponding c.p. modes on different instruments should correlate significantly and positively (P scores obtained by different instruments should show a significant positive correlation) and patterns of interrelationships between

preference modes should be similar within and across instruments (the from the literature well known Q with R correlation should be found between the Q and R scores of each instrument).

These requirements lead to convergent validity.

- 2) Different preference modes within and across different instruments (methods) should not correlate as strongly as corresponding traits (the correlation between R and P measured by the same or different instruments should be lower than the correlation between the P scores of two instruments).

This requirement leads to discriminant validity.

The c.p. inventory in this study provided ipsatized scores which are not independent mathematically and follow a sampling distribution with a mean of $-.33$ rather than the familiar sampling distribution of correlations which has a mean of 0 (Hicks, 1970). The correlations across instruments follow the familiar distribution with mean $"0"$.

This made assessment of discriminant validity difficult and prompted us to concentrate on convergent validity.

Note that the multitrait-multimethod methodology still confounds instrument validity with construct validity. For example low correlations of corresponding preference modes across instruments may be due to: a) lack of construct validity, b) poor instrument validity for one or more instruments, and c) plain lack of reliability on one or more of the instruments.

Instruments, Samples and Procedures

Participating students responded to two c.p. instruments. Their instructors rated a subsample of students on various c.p. related behaviors.

The Science Cognitive Preference Inventory (SCPI) is a 30 item instrument (Tamir et. al., 1976) which follows the conventional format of c.p. measurement described above. Test items cover physical and life science topics and presuppose a knowledge and understanding of science at the 9th or 10th grade level. The instrument had been extensively piloted and modified before use in this study. A full description of the SCPI, instructions for scoring, and data for various populations ranging from 9th grade to college may be found in Van den Berg (1978). Cronbach alpha reliability coefficients obtained in this study are shown in Table I.

The Reading Instrument (RI) was designed to measure student cognitive preferences in a different and more natural way than conventional c.p. inventories. In a science reading assignment covering four different topics in 15 paragraphs students are asked to indicate the most and least interesting sentences in each paragraph. Since each sentence is constructed such that it can be clearly classified in one of the c.p. modes (classification by 3 judges resulted in a better than 85% full agreement for each of the four modes), student responses may be scored using an algorithm described in Van den Berg (1978). A sample paragraph is the following (c.p. classification of each sentence indicated in parentheses):

Plants that carry out photosynthesis are called producers, other organisms including people) are called consumers (R). In the evolution producers must have originated earlier than consumers, since consumers cannot live without producers (P). It has been estimated that 85% of the photosynthesis on earth takes place in the oceans (R). From what source would ocean plants get the carbon that is necessary for photosynthesis? (Q). Some futurists expect that "ocean agriculture" will solve the world's food problems (A)

The RI differs from the SCPI in that students responding to it are engaged in a quite common and natural information processing activity (reading about science) while students taking the SCPI are responding

to a rather unusual and unfamiliar kind of test. Common to both tests is that student preferences for explicitly stated information are being measured as opposed to spontaneous student behavior.

In the Teacher Rating Instrument (TRI) teachers are asked to rate student behaviors thought to be related to cognitive preference on a 5-point Likert scale. Sample questions are: "This student is interested in applications of science" (A-mode behavior), "This student often asks critical questions about the validity of scientific information" (Q-mode), "This student likes to learn general rules and principles that can be applied to many scientific phenomena" (P-mode). The teacher indicates through ratings 1-5 whether (s)he agrees or disagrees with the statement concerning the student being rated. Teachers are also asked to record any particular comments they may have about the behavior of the student in science class. Each mode of c.p. is represented by 2 statements the ratings of which are added to define a score. The instrument differs from the SCPI and RI since it attempts to measure c.p. through observation of spontaneous behavior.

The main sample in this study consisted of 71 elementary education majors enrolled in a compulsory science laboratory course at a midwestern university. The study was replicated in a small group of 37 10th grade students enrolled in a compulsory life science course in a small rural high school in order to get some indication of the generalizability of the first study to other populations.

In the main sample the SCPI was administered during the first day of the science laboratory course, some weeks later (due to scheduling problems) followed by the RI (SCPI 79 students, RI 71 due to withdrawal). The TRI was administered to the instructors of the course near the end of the semester. Student c.p.'s should not have changed over the period between the administration of the instruments since c.p.'s are supposed to be rather stable over time (test-retest correlations for the SCPI

were greater than .70 over a 5 week period). However, if student c.p.'s would have changed during the period between the administration of the SCPI and RI, correlations between corresponding preference modes of both instruments would decrease, resulting in too conservative estimates for the convergent validity. Furthermore, the time passed between taking the SCPI and the RI reduced the possibility of interaction between the instruments.

After both testing sessions some students were interviewed informally to check for any unusual reactions to the instruments and the testing situation.

The SCPI was administered using rating procedures, students were instructed to rate the four options to each SCPI item on a 1-6 point scale according to their interest. The rating scores were required for a parallel study of rating versus ranking procedures, however, for the purpose of this study the ratings were converted into rankings ("ipsatized") by computer using an algorithm described in Van den Berg (1978). Ipsatization proved to be necessary to remove variance due to individual rating tendencies of students ("response set", some people tend to rate all statements high, others low) from the original c.p. scores.

An important première in this study was the use of an index of distinctness of c.p. to distinguish students with an outspoken c.p. from others who don't have a strong preference for one c.p. mode or another. One might expect that person's with a strong or distinct c.p. could be recognized in the discrepancies between their c.p. scores for the four modes while for others the 4-score c.p. pattern would not show large differences between the four scores. The index of distinctness called "F" indicated the probability that a person's c.p. pattern had originated by random response to SCPI items. If this probability was low, then the pattern was called distinct. F was calculated using an

analysis of variance in a "preference modes" by "test items" design (analogous to a treatments by subjects design, Lindquist, 1953) using the original ratings of SCPI response statements rather than the ipsatized ratings. For extensive information regarding this interesting variable the reader be referred to Van den Berg (1978) and personal contact with the authors.

The "F" was used to define a subsample of students with distinct preferences. Such students would be expected to show a higher correlation between their c.p. scores on different instruments, thus in such a subsample one would expect a higher convergent validity of the construct.

Teacher ratings were obtained on 29 out of 71 students, 14 of whom had been selected because of their distinct score patterns while the remaining 15 students had been selected randomly. The two instructors each rated their own students.

The SCPI and the RI were administered within one lesson period of a second semester, tenth grade, compulsory life science course. The opportunity to work with this small sample was used to pilot the questions of whether the results of the main experiment were generalizable to the tenth grade level and to truly ipsative (rather than ipsatized) scores.

Results

The SCPI Cronbach alpha reliability coefficients (Table I) were judged to be quite sufficient for this study. Differences between the reliabilities of rating scores and ipsatized ratings illustrated the data reduction which took place through removal of the variance due to student response set as described in the procedures section. As in other c.p. studies the Q-mode appeared the most reliable.

TABLE I

Cronbach Alpha Reliability Coefficients of the Science Cognitive Preference Inventory

Samples	Modes of Cognitive Preference				
	N	R	P	Q	A
Elem. Ed. Majors rating scores	79	.84	.80	.87	.77
ipsatized scores	79	.59	.60	.79	.48
Tenth Graders ipsative scores	37	.57	.67	.69	.52

TABLE II

Pearson Correlations of Science Cognitive Preference Inventory (Ipsatized) Scores with Reading Instrument Scores of Elementary Education Majors

Science Cognitive Preference Inventory (SCPI)					
		R	P	Q	A
Total Sample (N = 71)					
Reading Instrument (RI)	R	.19	.02	-.08	-.11
	P	.04	.21	-.22	.07
	Q	-.39 ^{a)}	-.12	.43 ^{b)}	-.11
	A	.23	-.05	-.28	.25
Subsample with Beyond Chance (p .05) Distinct c.p. Patterns (N = 36)					
Reading Instrument (RI)	R	.20	.02	-.07	-.12
	P	.16 ^{b)}	.26	-.36 ^{a)}	.12
	Q	-.52 ^{b)}	-.13	.57 ^{b)}	-.22 ^{a)}
	A	.34	-.08	-.39 ^{a)}	.37 ^{a)}

a) $p < .05$

b) $p < .01$

INSERT TABLES I AND II ABOUT HERE

The reliability coefficients for the RI (Reading Instrument) were estimated to be below .40. Since scores for each of the preference modes were based on only two items, reliabilities for the TRI (Teacher Rating Instrument) were low.

Table II shows the correlations between the SCPI and RI scores for the total sample of elementary education majors and for a subsample of students with distinct score patterns. Correlations between corresponding modes of c.p. generally were somewhat higher than those between non-corresponding modes. For the total sample only the Q_{SCPI} correlated significantly with its corresponding Q_{RI} ($r=.43$). The relationship between Q and R found within the SCPI ($r=-.69$ which still is a strongly negative correlation in a sampling distribution of correlations which has a mean of $-.33$ rather than "0", see comment page 7) reproduced across instruments ($R_{SCPI}:Q_{RI}$, $r=-.39$, Table II) and within the RI ($R_{RI}:Q_{RI}$, $r=-.57$).

The SCPI-RI correlation pattern was more pronounced in the subsample of students with distinct c.p. score patterns, distinct here meaning that the probability that these score patterns had originated through random response to the SCPI items was less than .05. Convergence was found for the Q and A modes of c.p. with correlations of .57 and .37 respectively. Significant correlations were also found between $Q_{RI}:R_{SCPI}$, $P_{RI}:Q_{SCPI}$, and $A_{RI}:Q_{SCPI}$. The negative correlation observed between $P_{RI}:Q_{SCPI}$ also showed within the SCPI ($P_{SCPI}:Q_{SCPI}$, $r=-.50$) and within the RI ($P_{RI}:Q_{RI}$, $r=-.40$). Strongly negative correlations between P and Q had also been found by Tamir and Kempa (1976) and Lunetta and Tamir (1978) in different samples and with a different c.p. inventory. The negative correlation between A and Q also showed

within the RI ($A_{RI}:Q_{RI}$, $r=-.29$) but not within the SCPI ($Q_{SCPI}:A_{SCPI}$, $r=-.32$) while $-.33$ is the mean and point of no relationship for intercorrelations between 4-choice ipsative scores. The literature also does not provide consistent patterns for Q:A relationships.

The small replication study with 10th graders showed results consistent with those of the study with elementary education majors, except for the A dimension. Note also the convergence of the P-mode in this sample (Table III) and reciprocity of the significant P:Q across instruments correlations ($-.36$ and $-.42$, while the correlation of P_{SCPI} with Q_{SCPI} was $-.50$ for this sample).

 INSERT TABLES III AND IV ABOUT HERE

The TRI R and Q scores (Table IV) showed significant correlations with the SCPI R score, which showed consistency with the familiar negative Q:R correlations found in other studies. Strong correlations of the SCPI A score with the TRI R and Q scores were puzzling. The TRI may have some potential on the R dimension, however, for the other modes much improvement is needed.

Discussion and conclusions

The weaknesses of this study, i.e. the yet weak reliabilities of the Teacher Rating Instrument and the Reading Instrument, and the time lapse between the administration of the RI and SCPI, would be expected to introduce extra error variance and thus depress the between instrument correlations and result in conservative decisions regarding the convergent validity of the c.p. construct.

Strong evidence of convergent validity was obtained for the Q-mode

TABLE III

Pearson Correlations of Science Cognitive Preference Inventory
(Ipsative) Scores with Reading Instrument Scores of
Tenth Graders (N = 37)

Science Cognitive Preference Inventory (SCPI)					
		R	P	Q	A
Reading Instrument (RI)	R	.17	.15 ^{b)}	-.28 ^{a)}	-.13
	P	-.10	.44 ^{b)}	-.36 ^{a)}	-.02
	Q	-.16	-.42 ^{a)}	.48	.12
	A	-.05	-.02	.01	.19

a) $p < .05$

b) $p < .01$

TABLE IV

Kendall Tau Correlations of Science Cognitive Preference
Inventory (Ipsatized) Scores with Teacher Rating Instrument
Scores for a Subsample of Elementary Education Majors (N = 29)

Science Cognitive Preference Inventory (SCPI)					
		R	P	Q	A
Teacher Rating Instrument (TRI)	R	.22 ^{a)}	-.12	-.19	-.38 ^{a)}
	P	.11	-.01	-.13	-.10
	Q	-.22 ^{a)}	.10	-.18	.34 ^{a)}
	A	-.03	-.03	.06	-.09

a) $p < .05$

of c.p. (both samples) while the A-mode (only on the RI of the main sample), the P-mode (on the RI, 10th grade sample) and R-mode (TRI, main sample) showed a much weaker convergent validity.

SCPI intercorrelation patterns for P and Q, and R and Q were duplicated within the Reading Instrument and between the RI and SCPI, adding to the validity of the P, Q, and R-modes of c.p.

The major results were replicated in a small pilot sample of 10th graders who had taken the SCPI using ranking rather than rating.

The correlations found were moderate, but acceptable in this kind of research (see examples in Campbell and Fiske, 1959). Furthermore cognitive style measures have been rather sensitive to changes in measurement method. For example, correlations between the well established Embedded Figures Test of field-dependent..independent cognitive style with three other measures of field-dependence have been reported to be .82, .71, and .39 for males, and .63, .55 and .34 for females, the latter two figures being for measures thought to have better construct validity (ABC degree of body articulation and Portable Rods Frame Task, Oltman et. al., 1971). Gray (1974) reported convergency correlations of .47, .52, and .65 between two tests of categorical-descriptive-relational cognitive style, the tests differed in that one test used photographs and the other line drawings for the pictures to be shown.

Calling the results of this study acceptable in the light of results of comparable studies does not eliminate all doubts regarding the validity and utility of cognitive styles including cognitive preferences, which are supposed to be relatively stable modes of conceptualization of the environment. Sensitivity for methods would suggest some lack of generalizability of cognitive styles across different kinds of environmental stimuli (methods).

The index of distinctness indicated that 49% of the elementary education majors showed c.p. score patterns which could have originated by chance. Such people cannot be said to have a cognitive preference that is measurable with a c.p. inventory. Studies have shown that the percent of people with a distinct c.p. varies from 20% in a 9th grade sample to 78% in a sample of highly motivated high school students (Van den Berg, 1978). In this study distinct preferences showed more stability across across methods than non-distinct preferences. Jungwirth (1978) showed that distinct preferences are more stable with respect to time as well. The distinctness appeared to be a very relevant characteristic of c.p. which could be very helpful in future studies. The index could assist researchers in selecting samples of people with distinct or even extreme preferences for special study, however, the index could also be used as a variable in its own right.¹

The c.p. construct and c.p. inventories had been invented to assess preferences for types of intellectual responses to information. However, present c.p. inventories only assess preferences for inquiry and non-inquiry-type statements but not actual inquiry behavior exhibited by students. Unless relationships can be established between c.p. and such behaviors, the observed shift of student preferences toward P and Q observed in science curriculum evaluation studies (Tamir, 1977) does not have a clear meaning.

1) Preliminary copies of a publication with extensive information regarding the distinctness variable will be available shortly from the authors.

The Teacher Rating Instrument (TRI) was not successful at assessing such relationships but the instrument may be improved. For example, in this study the instructors did not see the TRI form until after the semester. In the future a more extensive instrument with more low inference items may be compiled and provided to the teachers during the semester to focus their attention on specific student behaviors to be observed.

The Reading Instrument (RI) still only measures student c.p. for explicit statements, however, the instrument could serve to link c.p. for explicit statements with relevant cognitive activities. In a follow-up retention test one could investigate differential retention as function of c.p., for example, would high Q students remember Q statements from the RI better than low Q students ? Furthermore, one could also ask students to summarize the information contained in the RI, for example to see whether high P students use more generalizations in summarizing than low P or high R students. The reliability of the RI could be increased by expanding the instrument (it presently only takes 10-15 minutes to administer) and by decreasing the number of different science topics relative to its length.

In conclusion it might be said that this study provided useful information regarding the validity of the c.p. construct. However, more studies are needed to establish the validity of commonly used interpretation of the c.p. construct. Future studies should first of all focus on the relationships of the c.p. construct and its measurement with actual student tendencies to respond to scientific information. Fringe benefits of such studies may be improvements of methods for the assessment of inquiry attitudes and behavior in science education.

References

- Brown, S.A. Cognitive preferences in science: Their nature and analysis. Studies in Science Education, 1975, 2, 43-65.
- Campbell, D.T., Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- Cronbach, L.J. Test validation. in: Thronthike, R.J. (ed.), Educational Measurement, (2nd edition), Washington D.C.: American Council on Education, 1971.
- Cronbach, L.J., Meehl, P.E. Construct validity in psychological tests. Psychological Bulletin, 1955, 52(4), 281-302.
- Gray, J.L. Stability and generalizability of conceptual style. Psychology in the Schools, 1974, 11(4), 466-475.
- Heath, R.H. Curriculum, cognition, and educational measurement. Educational and Psychological Measurement, 1964, 24, 239-253.
- Hicks, L.E. Some properties of ipsative, normative, and forced-choice normative measures. Psychological Bulletin, 1970, 74, 167-184.
- Jungwirth, E. Cognitive consistency in cognitive preference tests and its implications for test-validity. Studies in Educational Evaluation, 1978, 4(2).
- Jungwirth, E. Cognitive preference testing in the natural sciences: Some question marks. European Journal of Science Education, 1979a, 1(4).
- Jungwirth, E. Cognitive preference testing in evaluation of curriculum implementation: A critical view. Unpublished paper, The Hebrew University of Jerusalem, P.O. Box 12, Rehovot, Israel, 1979b.
- Jungwirth, E. An alternative interpretation of findings in cognitive preference research in science education. Science Education, 1980, 64(1), 85-94.
- Lindquist, E.F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, 1953.
- Tamir, P. The relationship among cognitive preference, school environment, teacher's curricular bias, curriculum and subject matter. American Educational Research Journal, 1975, 12, 235-264.
- Tamir, P. A note on cognitive preferences in science. Studies in Science Education, 1977, 4, 111-121.
- Tamir, P. et. al. Science Cognitive Preference Inventory. Curriculum, Research, and Development Group, University of Hawaii, 1976.
- Tamir, P., Kempa, R.F. College students' cognitive preferences in science. The Journal of Educational Research, 1976/77, 70, 210-218.
- Tamir, P., Lunetta, V.N. A comparison of ipsative and normative procedures in the study of cognitive preferences. Paper presented at the National Annual Meeting of the American Educational Research Association, New York, April 1977.
- Van den Berg, E. Cognitive Preferences: A Validation Study. Unpublished doctoral dissertation, The University of Iowa, 1978.
- Williams, C. A study of cognitive preferences. Journal of Experimental Education, 1975, 43, 61-77.